Self-normalized deviation inequalities with application to t-statistic

Xiequan Fan*

Center for Applied Mathematics, Tianjin University, 300072 Tianjin, China

Abstract

Let $(\xi_i)_{i\geq 1}$ be a sequence of independent and symmetric random variables. We obtain some upper bounds on tail probabilities of self-normalized deviations

$$\mathbf{P}\Big(\max_{1\leq k\leq n} \sum_{i=1}^{k} \xi_i \Big/ \big(\sum_{i=1}^{n} |\xi_i|^{\beta} \big)^{1/\beta} \geq x \Big)$$

for $x > 0$ and $\beta > 1$. Our bound is the best that can be obtained from the Bernstein inequality under the present assumption. An application to Student's t-statistic is also given.

Keywords: Self-normalized deviations; Student's t-statistic; exponential inequalities

2000 MSC: Primary 60E15; 60F10

1. Introduction

Let $(\xi_i)_{i\geq 1}$ be a sequence of independent, centered and nondegenerate real-valued random variables (r.v.s). Denote by

$$S_n = \sum_{i=1}^{n} \xi_i \quad \text{ and } \quad V_n(\beta) = \Big(\sum_{i=1}^{n} |\xi_i|^{\beta} \Big)^{1/\beta}, \ \beta > 1.$$

The study of the tail probabilities $\mathbf{P}(S_n/V_n(\beta) \geq x)$ certainly has attracted some particular attentions. In the case where r.v.s $(\xi_i)_{i\geq 1}$ are identically distributed and $\mathbf{E}|\xi_1|^{\beta} = \infty, \beta > 1$, Shao [9] proved the following deep large deviation principle (LDP) result: for any $x > 0$,

$$\lim_{n\to\infty} \mathbf{P}\Big(\frac{S_n}{V_n(\beta)\, n^{1-1/\beta}} \geq x \Big)^{1/n} = \sup_{c\geq 0} \inf_{t\geq 0} \mathbf{E}\Big[\exp\Big\{ t\Big(cX - x\Big(\frac{1}{\beta}|X|^{\beta} + \frac{\beta-1}{\beta} c^{\beta/(\beta-1)} \Big) \Big) \Big\} \Big].$$

The related moderate deviation principles (MDP) are also given by Shao [9] and Jing, Liang and Zhou [6]. However, the LDP and MDP results do not diminish the need for tail probability inequalities valid for given n. Such inequalities have been obtained in particular by Wang and Jing [10]. For instance, they proved that if the r.v.s $(\xi_i)_{i\geq 1}$ are symmetric (around 0), then for all $x > 0$,

$$\mathbf{P}\Big(\frac{S_n}{V_n(2)} \geq x \Big) \leq \exp\Big\{ -\frac{x^2}{2} \Big\}. \tag{1}$$

*Corresponding author.

E-mail: fanxiequan@hotmail.com (X. Fan).

This bound is rather tight for moderate x's. Indeed, as showed by the MDP result of Shao [9] (cf. Theorem 3.1), for certain class of r.v.s it holds that

$$x_n^{-2} \ln \mathbf{P}\bigg(\frac{S_n}{V_n(2)} \geq x_n \bigg) = -\frac{1}{2}, \tag{2}$$

for $x_n \to \infty$ and $x_n = o(\sqrt{n}\,)$ as $n \to \infty$. See also Theorem 2.1 of Jing, Liang and Zhou [6] for non identically distributed r.v.s. In Fan, Grama and Liu [3], inequality (1) has been further extended to the case of partial maximum: for all $x > 0$,

$$\mathbf{P}\bigg(\max_{1 \leq k \leq n} \frac{S_k}{V_n(2)} \geq x \bigg) \leq \exp\left\{ -\frac{x^2}{2} \right\}. \tag{3}$$

On the other hand, by the Cauchy-Schwarz inequality, it is easy to see that $S_n^2 \leq n\,(V_n(2))^2$. Therefore, for all $x > \sqrt{n}$,

$$\mathbf{P}\bigg(\frac{S_n}{V_n(2)} \geq x \bigg) \leq \mathbf{P}\bigg(\max_{1 \leq k \leq n} \frac{S_k}{V_n(2)} \geq x \bigg) = 0, \tag{4}$$

which cannot be deduced from (1) and (3). Hence, the inequalities (1) and (3) are not tight enough for large x's.

In this paper we give an improvement on inequality (3); see inequality (5). Our inequality coincides with (4). More general, we establish an upper bound on tail probabilities $\mathbf{P}(\max_{1 \leq k \leq n} S_k / V_n(\beta) \geq x), x > 0$, for symmetric r.v.s $(\xi_i)_{i \geq 1}$. In particular, we show that our inequality is the best that can be obtained from the classical Bernstein inequality: $\mathbf{P}(X > x) \leq \inf_{\lambda > 0} \mathbf{E}[e^{\lambda(X - x)}]$. An application to Student's t-statistic is also given.

The paper is organized as follows. Our main results and applications are stated and discussed in Section 2. Proofs are deferred to Section 3.

2. Main results

In the following theorem, we give a self-normalized deviation inequality for independent and symmetric random variables.

Theorem 2.1. *Assume that $(\xi_i)_{i \geq 1}$ is a sequence of independent, symmetric and nondegenerate random variables. Denote by*

$$V_n(\beta) = \Big(\sum_{i=1}^n |\xi_i|^\beta \Big)^{1/\beta}, \qquad \beta \in (1, \infty).$$

Then for all $x > 0$,

$$\mathbf{P}\left(\max_{1 \leq k \leq n} \frac{S_k}{V_n(\beta)} \geq x \right) \leq B_n(\beta, x) := \frac{1}{2^n} \left(\sqrt{t} + \frac{1}{\sqrt{t}} \right)^n t^{-\frac{1}{2} n^{1/\beta} x} \, \mathbf{1}_{\{x \leq n^{(\beta-1)/\beta}\}}, \tag{5}$$

where

$$t = \frac{n^{(\beta-1)/\beta} + x}{n^{(\beta-1)/\beta} - x}$$

with the convention that $B_n(\beta, n^{(\beta-1)/\beta}) = 2^{-n}$. *Moreover,* $B_n(2,x)$ *is increasing in* n *and for any* $x > 0$,

$$\lim_{n\to\infty} B_n(2,x) = \exp\left\{-\frac{x^2}{2}\right\}. \tag{6}$$

Hölder's inequality implies that $S_n \le V_n(\beta) n^{(\beta-1)/\beta}$. Thus when $x > n^{(\beta-1)/\beta}$, it holds that

$$\mathbf{P}\Big(\max_{1\le k\le n} \frac{S_k}{V_n(\beta)} \ge x\Big) = 0. \tag{7}$$

This feature coincides with the fact that $B_n(\beta,x) = 0$ for all $x > n^{(\beta-1)/\beta}$.

Notice that bound (5) is the best that can be obtained from the following Bernstein inequality

$$\mathbf{P}\left(\frac{S_n}{V_n(\beta)} \ge x\right) \le \inf_{\lambda\ge 0} \mathbf{E}\Big[e^{\lambda\big(\frac{S_n}{V_n(\beta)} - x\big)}\Big]. \tag{8}$$

Indeed, if $\xi_i = \pm a, a > 0$, with probabilities 1/2, then it holds for all $0 < x < \sqrt{n}$,

$$\inf_{\lambda\ge 0} \mathbf{E}\Big[e^{\lambda\big(\frac{S_n}{V_n(\beta)} - x\big)}\Big] = \inf_{\lambda\ge 0} \mathbf{E}\Big[e^{\lambda\big(\frac{S_n}{an^{1/\beta}} - x\big)}\Big] = \inf_{\lambda\ge 0} e^{-\lambda x}\Big(\cosh(\frac{\lambda}{n^{1/\beta}})\Big)^n = B_n(\beta,x).$$

Moreover, when $x \nearrow n^{(\beta-1)/\beta}$, bound (5) tends to 2^{-n}, which is the best possible at $x = n^{(\beta-1)/\beta}$. Indeed, for the ξ_i's mentioned above, it holds

$$\mathbf{P}\left(\max_{1\le k\le n} \frac{S_k}{V_n(\beta)} \ge n^{(\beta-1)/\beta}\right) = \mathbf{P}\Big(\xi_i = a \text{ for all } i \in [1,n]\Big) = \frac{1}{2^n}.$$

By (6), it holds

$$\lim_{n\to\infty} B_n(2,x) = e^{-x^2/2} \to e^{-n/2}, \qquad x \to n^{1/2}, \tag{9}$$

which seems to be contracted with $B_n(2,x) \to 2^{-n}$ as $x \to n^{1/2}$. In fact, (6) holds for any fixed $x > 0$, but it doesn't hold for $x = c\,n^{1/2}$, where c is a positive absolute constant. For example, consider the following function

$$B_n(2,x) = \exp\Big\{-\frac{x^2}{2} - (\ln 2 - \frac{1}{2})\frac{x^4}{n}\Big\}, \quad x > 0.$$

Then the last function satisfies (9) and $\lim_{x\to n^{1/2}} B_n(2,x) = 2^{-n}$.

Since the r.v.s $(\xi_i)_{i\ge 1}$ are symmetric, it is obvious that for all $x > 0$,

$$\mathbf{P}\left(\max_{1\le k\le n} \frac{S_k}{V_n(\beta)} \le -x\right) \le B_n(\beta,x),$$

where $B_n(\beta,x)$ is defined by (5).

When $\beta \in (1,2]$, inequality (5) implies the following bound.

Corollary 2.1. *Assume condition of Theorem 2.1. If $\beta \in (1,2]$, then for all $x>0$,*

$$\mathbf{P}\bigg(\max_{1\le k\le n} \frac{S_k}{V_n(\beta)} \ge x\bigg) \le \exp\left\{-\frac{x^2}{2} n^{\frac{2}{\beta}-1}\right\}. \tag{10}$$

In particular, the last inequality implies that for any $\beta \in (1,2)$,

$$\frac{S_n}{V_n(\beta)} \to 0, \qquad n\to\infty,$$

in probability.

For $\beta \in (1,2]$, inequality (10) implies the following upper bound of LDP:

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbf{P}\bigg(\max_{1\le k\le n} \frac{S_k}{V_n(\beta) n^{(\beta-1)/\beta}} \ge x\bigg) \le -\frac{x^2}{2}, \qquad x\in(0,1]. \tag{11}$$

It also implies the following upper bound of MDP: for any $\alpha \in (\frac{\beta-2}{2\beta}, \frac{\beta-1}{\beta})$,

$$\limsup_{n\to\infty} \frac{1}{n^{2\alpha+\frac{2}{\beta}-1}} \ln \mathbf{P}\left(\max_{1\le k\le n} \frac{S_k}{V_n(\beta) n^{\alpha}} \ge x\right) \le -\frac{x^2}{2}, \qquad x\in(0,\infty). \tag{12}$$

With certain regularity conditions on tail probabilities of ξ_i, the LDP and MDP results are allowed to be established. We refer to Shao [9] and Jing, Liang and Zhou [6].

Wang and Jing [10] proved that for all $x>0$,

$$\mathbf{P}\bigg(\frac{S_n}{V_n(2)} \ge x\bigg) \le \exp\left\{-\frac{x^2}{2}\right\}. \tag{13}$$

An earlier result similar to (13) can be found in [5], where Hitczenko has obtained the same upper bound on tail probabilities $\mathbf{P}(S_n \ge x\, ||V_n(2)||_\infty)$. When $\beta=2$, inequality (10) reduces to the following inequality of Fan *et al.* [3]: for all $x>0$,

$$\mathbf{P}\bigg(\max_{1\le k\le n} \frac{S_k}{V_n(2)} \ge x\bigg) \le \exp\left\{-\frac{x^2}{2}\right\}. \tag{14}$$

Thus inequality (10) can be regarded as a generalization of (13) and (14). Moreover, since bound (5) is less than bound (14), our inequality (5) improves on (14).

If $(\xi_i)_{i\ge1}$ have $(2+\delta)$th moments with $0<\delta\le1$, the inequalities (13) and (14) can be further improved. For instance, Jing, Shao and Wang [7] proved the following Cramér type large deviations for i.i.d. (not necessarily symmetric) r.v.s:

$$\mathbf{P}\Big(S_n \ge xV_n(2)\Big) = \Big(1-\Phi(x)\Big)\Big(1+o(1)\Big), \quad n\to\infty, \tag{15}$$

uniformly for all $0\le x = o(n^{\delta/(4+2\delta)})$. Similarly, Liu, Shao and Wang [8] proved the following result for the maximum of sums:

$$\mathbf{P}\Big(\max_{1\le k\le n} S_k \ge xV_n(2)\Big) = 2\Big(1-\Phi(x)\Big)\Big(1+o(1)\Big), \quad n\to\infty, \tag{16}$$

uniformly for all $0 \leq x = o(n^{\delta/(4+2\delta)})$. Moreover, these asymptotic estimations are also more precise than (5) for moderate x's.

Let $(Y_i)_{i\geq 1}$ be a sequence of independent nondegenerate r.v.s, and $(d_i)_{i\geq 1}$ be a sequence of independent Rademacher r.v.s, i.e. $\mathbf{P}(d_i = \pm 1) = \frac{1}{2}$. Let $\xi_i = d_i Y_i$. Assume that $(Y_i)_{i\geq 1}$ and $(d_i)_{i\geq 1}$ are independent. Then we now have

$$S_n = \sum_{i=1}^{n} d_i Y_i, \qquad V_n(\beta) = \Big(\sum_{i=1}^{n} |Y_i|^{\beta}\Big)^{1/\beta}, \qquad \beta > 1.$$

The following result easily follows from Theorem 2.1 and Corollary 2.1.

Corollary 2.2. *Let $\xi_i = d_i Y_i$ for all $i \geq 1$. If $\beta \in (1, 2]$, then for all $0 < x \leq n^{(\beta-1)/\beta}$,*

$$\begin{aligned}\mathbf{P}\left(\max_{1\leq k\leq n} \frac{S_k}{V_n(\beta)} \geq x\right) &\leq B_n(\beta, x) \\ &\leq \exp\left\{-\frac{x^2}{2} n^{\frac{2}{\beta}-1}\right\}.\end{aligned}$$

In particular, the last inequality implies that for any $\beta \in (1, 2)$,

$$\frac{S_n}{V_n(\beta)} \to 0, \qquad n \to \infty,$$

in probability.

Consider Student's t-statistic T_n defined by

$$T_n = \sqrt{n}\,\overline{\xi}_n/\widehat{\sigma},$$

where

$$\overline{\xi}_n = \frac{S_n}{n} \quad \text{and} \quad \widehat{\sigma}^2 = \sum_{i=1}^{n} \frac{(\xi_i - \overline{\xi}_n)^2}{n-1}.$$

It is known that for all $x > 0$,

$$\mathbf{P}\Big(T_n \geq x\Big) = \mathbf{P}\Big(\frac{S_n}{\sqrt{[S]_n}} \geq x\Big(\frac{n}{n + x^2 - 1}\Big)^{1/2}\Big);$$

see Efron [1]. Notice that for all $x > 0$, it holds that $0 < x\Big(\frac{n}{n+x^2-1}\Big)^{1/2} < n^{1/2}$. Using inequality (5), we have the following exponential bound for Student's t-statistic.

Theorem 2.2. *Assume that $(\xi_i)_{i\geq 1}$ is a sequence of independent, symmetric and nondegenerate random variables. Then for all $x > 0$,*

$$\mathbf{P}\Big(T_n \geq x\Big) \leq B_n\Big(2,\ x\Big(\frac{n}{n + x^2 - 1}\Big)^{1/2}\Big), \tag{17}$$

where $B_n(2, x)$ is defined by (5).

3. Proofs of Theorem 2.1 and Corollary 2.1

The proof of Theorem 2.1 is based on a method called change of probability measure for martingales. The method is developed by Grama and Haeusler [4]. See also Fan, Grama and Liu [2].

Proof of Theorem 2.1. For any $i \geq 1$, set

$$\eta_i = \frac{\xi_i}{V_n(\beta)}, \quad \mathcal{F}_0 = \sigma\Big(|\xi_j|, 1 \leq j \leq n\Big) \text{ and } \mathcal{F}_i = \sigma\Big(\xi_k, 1 \leq k \leq i,\ |\xi_j|, 1 \leq j \leq n\Big). \tag{18}$$

Since $(\xi_i)_{i\geq 1}$ are independent and symmetric, then

$$\mathbf{E}[\xi_i > y \,|\, \mathcal{F}_{i-1}] = \mathbf{E}\Big[\xi_i > y \,\Big|\, |\xi_i|\Big] = \mathbf{E}\Big[-\xi_i > y \,\Big|\, |-\xi_i|\Big] = \mathbf{E}[-\xi_i > y \,|\, \mathcal{F}_{i-1}].$$

Thus $(\eta_i, \mathcal{F}_i)_{i=1,...,n}$ is a sequence of conditionally symmetric martingale differences, i.e. $\mathbf{E}[\eta_i > y|\, \mathcal{F}_{i-1}] = \mathbf{E}[-\eta_i > y|\, \mathcal{F}_{i-1}]$. It is easy to see that

$$\frac{S_n}{\mathrm{V}_n(\beta)} = \sum_{i=1}^{n} \eta_i \tag{19}$$

is a sum of martingale differences, and that $(\eta_i, \mathcal{F}_i)_{i=1,...,n}$ satisfies

$$\sum_{i=1}^{n} |\eta_i|^\beta = \sum_{i=1}^{n} \frac{|\xi_i|^\beta}{V_n(\beta)^\beta} = 1.$$

For any $x > 0$, define the stopping time T:

$$T(x) = \min\Big\{k \in [1, n] : \sum_{i=1}^{k} \eta_i \geq x\Big\},$$

with the convention that $\min \emptyset = 0$. Then it follows that

$$\mathbf{1}_{\{\max_{1\leq k\leq n} S_k/V_n(\beta) \geq x\}} = \sum_{k=1}^{n} \mathbf{1}_{\{T(x)=k\}}.$$

For any nonnegative number λ, define the martingale $M(\lambda) = (M_k(\lambda), \mathcal{F}_k)_{k=0,...,n}$, where

$$M_k(\lambda) = \prod_{i=1}^{k} \frac{\exp\{\lambda\eta_i\}}{\mathbf{E}\left[\exp\{\lambda\eta_i\}\,|\mathcal{F}_{i-1}\right]}, \qquad M_0(\lambda) = 1.$$

Since T is a stopping time, then $M_{T\wedge k}(\lambda)$, $\lambda > 0$, is also a martingale. Define the conjugate probability measure $\mathbf{P}_\lambda$ on $(\Omega, \mathcal{F})$:

$$d\mathbf{P}_\lambda = M_{T\wedge n}(\lambda)d\mathbf{P}. \tag{20}$$

Denote by $\mathbf{E}_\lambda$ the expectation with respect to $\mathbf{P}_\lambda$. Using the change of probability measure (20), we have for all $x > 0$,

$$\begin{aligned}\mathbf{P}\Big(\max_{1\leq k\leq n} \frac{S_k}{V_n(\beta)} \geq x\Big) &= \mathbf{E}_\lambda\left[M_{T\wedge n}(\lambda)^{-1}\mathbf{1}_{\{\max_{1\leq k\leq n} S_k/V_n(\beta) \geq x\}}\right] \\ &= \sum_{k=1}^{n} \mathbf{E}_\lambda\Big[\exp\Big\{-\lambda\sum_{i=1}^{k}\eta_i + \Psi_k(\lambda)\Big\}\mathbf{1}_{\{T(x)=k\}}\Big],\end{aligned} \tag{21}$$

where

$$\Psi_k(\lambda) = \sum_{i=1}^{k} \log \mathbf{E}\left[\exp\{\lambda\eta_i\} \Big| \mathcal{F}_{i-1}\right].$$

Since $(\eta_i, \mathcal{F}_i)_{i=1,\dots,n}$ is conditionally symmetric, one has

$$\mathbf{E}\left[\exp\{\lambda\eta_i\} | \mathcal{F}_{i-1}\right] = \mathbf{E}\left[\exp\{-\lambda\eta_i\} | \mathcal{F}_{i-1}\right],$$

and thus it holds

$$\mathbf{E}\left[\exp\{\lambda\eta_i\} | \mathcal{F}_{i-1}\right] = \mathbf{E}\left[\cosh(\lambda\eta_i) | \mathcal{F}_{i-1}\right]. \tag{22}$$

Since

$$\cosh(x) = \sum_{k=0}^{\infty} \frac{1}{(2k)!} x^{2k}$$

is an even function, then $\cosh(\lambda\eta_i)$ is $\mathcal{F}_{i-1}-$measurable. Thus (22) implies that

$$\mathbf{E}[\exp\{\lambda\eta_i\} | \mathcal{F}_{i-1}] = \cosh(\lambda\eta_i).$$

Notice that the function $g(x) = \log(\cosh(x))$ is even and convex in $x \in \mathbf{R}$ and increasing in $x \in [0, \infty)$. Since $|\sum_{i=1}^n \eta_i| \le n^{1-1/\beta}(\sum_{i=1}^n |\eta_i|^\beta)^{1/\beta} = n^{1-1/\beta}$, it holds

$$\Psi_k(\lambda) \le \Psi_n(\lambda) = \sum_{i=1}^{n} g\Big(\lambda\eta_i\Big) \le ng\Big(\frac{1}{n}\sum_{i=1}^{n} \lambda\eta_i\Big) \le ng\Big(\frac{\lambda}{n^{1/\beta}}\Big).$$

By the fact $\sum_{i=1}^k \eta_i \ge x$ on the set $\{T(x) = k\}$, inequality (21) implies that for all $x > 0$,

$$\begin{aligned}
\mathbf{P}\bigg(\max_{1\le k\le n} \frac{S_k}{V_n(\beta)} \ge x\bigg) &\le \sum_{k=1}^{n} \mathbf{E}_\lambda\left[\exp\left\{-\lambda x + ng\Big(\frac{\lambda}{n^{1/\beta}}\Big)\right\} \mathbf{1}_{\{T=k\}}\right] \\
&\le \exp\left\{-\lambda x + ng\Big(\frac{\lambda}{n^{1/\beta}}\Big)\right\} \mathbf{E}_\lambda\left[\sum_{k=1}^{n} \mathbf{1}_{\{T=k\}}\right] \\
&\le \exp\left\{-\lambda x + ng\Big(\frac{\lambda}{n^{1/\beta}}\Big)\right\}.
\end{aligned} \tag{23}$$

The last inequality attains its minimum at

$$\lambda = \lambda(x) = \frac{n^{1/\beta}}{2} \log\left(\frac{n^{(\beta-1)/\beta} + x}{n^{(\beta-1)/\beta} - x}\right), \qquad x \in (0, n^{(\beta-1)/\beta}).$$

Substituting $\lambda = \lambda(x)$ in (23), we obtain the desired inequality (5) for all $x \in (0, n^{(\beta-1)/\beta})$. When $x = n^{(\beta-1)/\beta}$, we have

$$\mathbf{P}\bigg(\max_{1\le k\le n} \frac{S_k}{V_n(\beta)} \ge n^{(\beta-1)/\beta}\bigg) = \lim_{x\nearrow n^{(\beta-1)/\beta}} \mathbf{P}\bigg(\max_{1\le k\le n} \frac{S_k}{V_n(\beta)} \ge x\bigg) \le \lim_{x\nearrow n^{(\beta-1)/\beta}} B_n(\beta, x) = 2^{-n}.$$

When $x > n^{(\beta-1)/\beta}$, the desired inequality follows from (7).

Notice that the function $h(x) = g(\sqrt{x})$ is convex and increasing in $x \in [0, \infty)$. Therefor $g(\sqrt{x})/x$ is increasing in x, and $g(\sqrt{\lambda^2/n})/(\lambda^2/n)$ is decreasing in n. Thus

$$B_n(2, x) = \inf_{\lambda \geq 0} \exp\left\{ -\lambda x + ng\Big(\frac{\lambda}{n^{1/2}}\Big) \right\}$$

is increasing in n. Since $ng(\frac{\lambda}{n^{1/2}}) \to \lambda^2/2, n \to \infty$, we obtain

$$\lim_{n\to\infty} B_n(2, x) = \sup_n B_n(2, x) = \inf_{\lambda \geq 0} \exp\left\{ -\lambda x + \frac{\lambda^2}{2} \right\} = \exp\left\{ -\frac{x^2}{2} \right\}.$$

This completes the proof of Theorem 2.1. □

Proof of Corollary 2.1. Since $\cosh(x) \leq \exp\{x^2/2\}$, we have

$$ng\Big(\frac{\lambda}{n^{1/\beta}}\Big) \ \leq \ \frac{\lambda^2}{2} n^{1-\frac{2}{\beta}} \tag{24}$$

for all $\lambda > 0$. Thus, from (23), for all $x > 0$,

$$\begin{aligned} \mathbf{P}\bigg(\max_{1\leq k\leq n} \frac{S_k}{V_n(\beta)} \geq x \bigg) \ &\leq \ \inf_{\lambda > 0} \exp\left\{ -\lambda x + ng\Big(\frac{\lambda}{n^{1/\beta}}\Big) \right\} \\ &\leq \ \inf_{\lambda > 0} \exp\left\{ -\lambda x + \frac{\lambda^2}{2} n^{1-\frac{2}{\beta}} \right\} \\ &= \ \exp\left\{ -\frac{x^2}{2}\, n^{\frac{2}{\beta}-1} \right\}, \end{aligned} \tag{25}$$

which gives the desired inequality (10). □

Acknowledgements

The author would like to thank the anonymous referee for valuable comments and remarks. The work has been partially supported by the National Natural Science Foundation of China (Grants no. 11601375 and no. 11626250).

References

[1] B. Efron, Student's t-test under symmetry conditions, J. Amer. Statist. Assoc. 64 (328)(1969) 1278-1302.

[2] X. Fan, I. Grama, Q. Liu, Hoeffding's inequality for supermartingales, Stochastic Process. Appl. 122 (2012) 3545–3559.

[3] X. Fan, I. Grama, Q. Liu, Exponential inequalities for martingales with applications, Electron. J. Probab. 20, no. 1 (2015) 1–22.

[4] I. Grama, E. Haeusler, Large deviations for martingales via Cramér's method, Stochastic Process. Appl. 85 (2000) 279–293.

[5] P. Hitczenko, Upper bounds for the L_p-norms of martingales, Probab. Theory Related Fields 86 (1990) 225–238.

[6] B.Y. Jing, H.Y. Liang, W. Zhou, Self-normalized moderate deviations for independent random variables, Sci. China Math. 55 (11) (2012) 2297–2315.

[7] B.Y. Jing, Q.M. Shao, Q.Y. Wang, Self-normalized Cramér-type large deviations for independent random variables. Ann. Probab. 31(4) (2003) 2167–2215.

[8] W. Liu, Q.M. Shao, Q.Y. Wang, Self-normalized Cramér type moderate deviations for the maximum of sums, Bernoulli 19(3) (2013) 1006–1027.

[9] Q.M. Shao, Self-normalized large deviations, Ann. Probab. 25 (1997) 285–328.

[10] Q.Y. Wang, B.Y. Jing, An exponential nonuniform Berry-Esseen bound for self-normalized sums, Ann. Probab. 27 (4) (1999) 2068–2088.